

DOCUMENT RESUME

ED 254 547

TM 850 156

TITLE Educational Evaluation and Assessment in the United States. Position Paper and Recommendations for Action.

INSTITUTION Council of Chief State School Officers, Washington, D.C.

PUB DATE 13 Nov 84

NOTE 13p.; Paper presented at the Annual Meeting of the Council of Chief State School Officers (Wilmington, DE, November 13, 1984).

AVAILABLE FROM Council of Chief State School Officers, 400 North Capitol Street, Suite 379, Washington, DC 20001.

PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01 Plus Postage. FC Not Available from EDRS.

DESCRIPTORS *Agency Role; Data Collection; Decision Making; *Educational Assessment; *Educational Improvement; Elementary Secondary Education; *Evaluation Needs; Government Role; Information Needs; Policy Formation; *Public Agencies

IDENTIFIERS Council of Chief State School Officers; Department of Education; Educational Indicators; *Evaluation Problems

ABSTRACT

The lack of high quality data to describe and monitor U.S. educational systems hampers reform decisions and will impair reform evaluation. This paper sets forth the Council of Chief State School Officers' (CCSSO) recommendations for the Council and federal, state, and local education agencies to help improve the quality of educational indicators. The CCSSO will assume a national leadership role in the formulation and coordination of educational assessment policies. Recommendations to the U.S. Department of Education include: (1) integrating evaluation project reports; (2) revising the "State Education Statistics" (popularly known as Secretary Bell's "Wall Chart"); (3) providing financial aids for state assessment programs; and (4) continuing joint federal/state/local planning. It is recommended that each state: (1) develop a systematic approach to monitoring education in relationship to state goals and context; (2) establish high standards for its collection and presentation of outcome indicators; and (3) work with other states to establish a core set of indicators for across-state comparisons. (BS)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED254547



Council of Chief State School Officers

Education Evaluation and Assessment

In the United States

Position Paper

and

Recommendations for Action

Education
...a sound
investment in
AMERICA.

"PERMISSION TO REPRODUCE THIS
MATERIAL IN MICROFICHE ONLY
HAS BEEN GRANTED BY

W.F. Pierce

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

This document has been reproduced as
received from the person or organization
originating it.
Minor changes have been made to improve
reproduction quality.

Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

EDUCATION EVALUATION AND ASSESSMENT IN THE UNITED STATES

Position Paper and Recommendations for Action

Adopted by the

Council of Chief State School Officers

at their

1984 Annual Meeting

November 13, 1984

Wilmington, Delaware

**Council of Chief State School Officers
400 North Capitol Street
Suite 379
Washington, D. C. 20001
(202) 393-8161**

EDUCATION EVALUATION AND ASSESSMENT IN THE UNITED STATES

Introduction

During the past two years there has been an intense and productive nationwide debate about the condition of education in the United States and how to improve it. The consequences of this debate have touched almost every community in the nation as the several states have carried out reviews of their policies and proposed significant new reforms. However, as citizens' groups, legislators, governors, and local and state boards of education have examined educational problems, they have often been frustrated by the lack of good information about the condition of education.

High quality data are essential to the continuing momentum of education reform. The publication of data brings education to the attention of the public and makes it an issue of enduring concern. By comparing the performance of a school, district, or state with itself over time, or with other schools, districts, or states, data serve to exhort, motivate, or reward.

A second use is to aid local, state, and national policymakers in understanding the consequences of changes in policy and to aid them in implementing policies once they are adopted. Because of the variety that exists across the nation's school districts, numerous natural experiments are continually taking place as academic standards are increased, steps are taken to enhance the quality of teaching, the school day or year is lengthened, or the level of resources available for educating specific groups of students changes. By observing the impact that these changes have on the outcomes of schooling, policymakers can estimate the costs and benefits of pursuing alternative strategies for improving student performance.

Finally, and coming full circle, data serve to make education accountable to the public. At no time have more citizens been involved at all levels, studying the problems of education and proposing solutions. In many states and localities, significant new dollars have been made available to fund these proposals. With this commitment comes a concomitant interest in the assessment and evaluation of the system and the systematic reporting of its strengths and weaknesses, successes, and failures.

Examples of the weakness of present educational information abound. Here are but two:

1. The National Science Foundation (NSF) reported a major national problem with the supply of qualified high school teachers of mathematics and science. Shortly afterward, the General Accounting Office found that data supporting NSF's conclusion were too weak as a basis for reasonable policy; and the problem, if there was one, was limited to a few regions rather than the entire nation.

2. The U.S. Department of Education published a statistical chart in 1984 which contained the "most comprehensive data available on the performance of state educational systems." The purpose was to compare states so that state leaders could "better consider appropriate reforms."

BEST COPY

The only measures of student achievement available for all states were the college entrance examinations—the Scholastic Aptitude Test (SAT) and the American College Entrance Examination (ACT). These scores are not representative of the quality of the states' educational systems. They do not reflect the diversity of the students or programs. They do not assess the quality of the elementary or middle school. At best they represent an assessment of the aptitude of college going twelfth graders.

The lack of high quality data to describe and monitor educational systems in the U.S. hampers reform decisions and will impair attempts to evaluate the effects of the reforms. We need valid ways of assessing the quality of education on a continuing basis.

Steps are being taken to improve assessment and evaluation information to meet this problem. At the federal level, the U.S. Department of Education has asked the National Academy of Sciences to examine the federal systems of data collection and analysis. The Department has also established a Federal Indicators Project in collaboration with nationally based state organizations to define the key statistical indicators that need to be collected by the federal government.

These are important, but not sufficient, steps to deal with problems that affect the entire educational system, from the federal level to the classroom. The Council of Chief State School Officers has worked over the past year to develop a set of recommendations for the Council and federal, state, and local education agencies, designed to help improve the quality of educational indicators. This paper sets forth the recommendations.

National Concerns

Since the birth of the Office of Education in 1867, the collection and dissemination of national statistics on education has been a federal responsibility. The federal government uses statistics to assess progress toward federal goals such as the provision of opportunities for the handicapped, to track federal funds to ensure that their use meets the intent of Congress, to assess compliance with civil rights statutes, and to monitor the health of the educational system in general. This last purpose leads to the collection of large amounts of data describing the financing of the system and the quantity of resources used in the schools. Less complete data are gathered on the quality of the resources and the outcomes of the educational process. Yearly, the Department assembles a three foot stack of reports containing thousands of statistics. Although many of the federal reports are well done and useful to both policy makers and researchers, they do not provide a good base for policy making.

To some extent, this problem exists because a variety of different agencies within the federal government produce statistical reports, including agencies both within and outside of the Department of Education. Better coordination across agencies can address this aspect of the problem. To a greater extent, however, the problem exists because there is little or no organizing theory or model to guide the collection and presentation of educational data.

The Federal Indicators Project and the National Academy Study show promise for improving the quality and use of federally collected data on education. Equally important is the federal sponsorship of research efforts to improve the quality of educational assessment and evaluation. Of special importance is research work on the measurement of student achievement, both for use by teachers in the classroom and for policy makers interested in aggregate measures of achievement at the school, district, state, and federal levels.

The need for these efforts is strongly supported by even a cursory analysis of Secretary Bell's 1984 "wall chart" entitled "State Education Statistics." The chart contains only two student "outcome" measures (graduation rate and college entrance examination test score means) and each has very limited value. There is no measure of teacher quality or of the quality of the curriculum. There is no measure of the opportunity for students to take academic courses. The quality of the available data was influenced by the Secretary's decision to contrast states; this required that there must be common data available from each state. But even if the Secretary had focused only on nationwide data, the results would have been unsatisfactory. Of the problems listed above, only the limits of achievement test information would have been even partially overcome.

The publication of the "wall chart" also brought to the forefront the issue of state-to-state comparisons. For years, comparisons among states on a wide variety of education measures have existed within the pages of Department of Education reports. These reports have even included rankings of the states on such measures as per pupil expenditure and teacher/pupil ratio. National attention to the "wall chart" was attracted by two things: first, the Secretary held a press conference and presented the chart, which was perceived by the press as a "report card" for the states; and second, the chart highlighted student outcomes and ranked states on them, thereby giving the media of every state a local story.

On a technical level, state-to-state comparisons are problematical for a variety of reasons beyond those given above concerning the quality of the data. The most important reason is that states differ greatly in the characteristics of their populations, the levels of available resources, and the curricula taught in their schools. These differences are particularly critical when student outcomes are being compared. Moreover, state comparisons on student achievement tests invoke the spectre of a single national exam, a vision that challenges the U.S. tradition of state and local control over education. On a political level, however, the attention given to the Secretary's "wall chart" makes inevitable future state to state comparisons on outcome measures.

These are important matters for the future of our nation's educational system. The task now is to make sure that the data are as high quality as possible and given appropriate interpretation. This has both short- and long-term implications. In the short run, the Secretary's "wall chart" for 1985, if there is one, must be different from the 1984 version. Over the long-term, the Federal Indicators Project, in conjunction with state and local groups, should provide a structure to improve the quality of data critical to federal concerns. This effort should result in standardized and improved specifications for gathering and reporting much of these data.

Whether such an effort should include federal specification of a set of indicators that would be collected by every state is another matter. The federal government can collect data that represent the status of the nation on core educational issues without having the data be representative at the state level. Yet, if we accept the argument that state-to-state comparisons are bound to happen, it is important to ensure that the comparisons be as policy relevant and as valid and fair as possible. An alternative to federal specification would be for the states collectively to provide the leadership. This would offer balance to the potential centralization of policy implicit in the federal government's taking the initiative. For much of the data to be collected, the eventual difference would probably be small. It might even be small on the most sensitive of issues: the comparison of states on achievement test results. The process for determining the policies in a democratic and decentralized system, however, is critical to the success of the policies.

A final nationwide issue needs to be considered. The current debates about education make much of our nation's economic and educational position vis a vis other nations. The natural experiment made possible by the great variations among nations in educational policies indicates that we might learn much by examining the differences among nations. Yet remarkably little systematic data are gathered on which to make legitimate comparisons among nations. The single international organization established for this purpose (the International Educational Achievement Organization—IEA) has received only sporadic support from the U.S. Government. It is time to address this issue.

State and Local Concerns

Arguments about the need for the nation to have a valid set of indicators collected over time, also hold at the state level. These indicators may be somewhat different from those viewed as necessary by the federal government. Each state may value certain outcomes differently, and no group of analysts has the "right" model. These differences, however, do not diminish the need for a state to be clear about the outcomes important to it or about the way that it assesses the effectiveness of resource use within its boundaries.

This point is particularly important now as the several states embark on programs of school reform. The quality of a state's efforts to evaluate the progress of its reforms will depend upon the quality of its educational data base. A solid data base containing regularly gathered measures of resources and outcomes will provide a foundation for monitoring the reforms. In the short-run, for example, the effects of many reforms on student achievement outcomes may not be noticeable. Instead they may be seen in such areas as increased access to academic courses or to dropout prevention programs. States with measures only of student achievement will know little about these issues. Similarly, states with only narrow achievement measures, as in some minimum competency examinations, will not be able to detect the effects of reforms like increased academic requirements, which might have more impact on high achieving students.

Additionally, because comparisons of states will occur with increasing frequency, it is important for the states to build the capacity to learn from these comparisons. For a state considering whether to increase its high school graduation requirements, the experience of similar states that have recently carried out such a policy change could be very informative. For this purpose it would be particularly useful for each state to have a common core of data that would enable valid and useful comparisons and a capacity to connect state data to nationwide information. These data, of course, could be collected on samples of schools and students, reducing the extra burden on the system.

Finally, the process of assessing and monitoring educational progress is important at the local level. Measures of outcome should be clearly related to the purposes and goals of the system, and there should be better ways of assessing the quality of resources in local education agencies and schools. In these ways the federal, state, and local needs are similar. There is still another need at the local level, however, which cannot be met by measures of institutional effectiveness. This is the need for ways of diagnosing the needs and directing the education of individual students. The theory, methods of construction, and use of these measures are qualitatively different from that of the outcome or achievement measures used for evaluating federal, state, and local systems. More needs to be done in the development of these measures, for they may be the most important of all.

Recommendations

This paper raises critical issues regarding educational assessment and evaluation and leads to recommendations for the CCSSO. These recommendations are addressed to the CCSSO itself and to federal, state and local governments. They cover a broad range of assessment and evaluation issues with a particular focus on the role that the CCSSO might take in providing direction and leadership over the next several years.

At this time of widespread reform in the United States, the CCSSO has an important responsibility for leadership in the development, use, and interpretation of assessment and evaluation procedures and results.

Education, assessment and evaluation take place at different levels: the individual student, the classroom, school, school district, state, nation, and world. A major challenge is to design assessment and evaluation procedures which will allow for efficient use of data for decision making at each level and, where appropriate, for use at multiple levels. Beyond the longer-range concern of tracking the health of the U.S. educational system to provide data for informed policy making, the task of developing models and implementing testing and data collection systems is of great immediate concern to the CCSSO, for without adequate data we will not be able to monitor and assess the impact of the various reforms. The recommendations that follow are directed to the CCSSO, to the U.S. Department of Education (and those responsible for federal appropriations—the U.S. Congress and the President), to the states individually, and to local education agencies.

To the Council of Chief State School Officers

1. Assume a national leadership role in the formulation and coordination of policies regarding the assessment of the K-12 educational system. Work actively with federal and state agencies to improve their capacity to gather, analyze, and report on a variety of assessment matters.
2. The Committee on Coordinating Educational Information and Research (CEIR) of the CCSO should monitor the development of state and national policies. It might also develop model policies for states to consider.
3. Establish in the CCSO a coordination center on assessment and evaluation programs of the states. The center would provide states with timely information on practices in other states; support cooperative efforts among states to align their programs more closely with one another; and promote coordination among international, national, and state assessment programs. The center shall be funded by the Council and/or outside sources. The scope of the work for the center and appropriate work schedule shall be prepared by the CEIR Committee and presented to the Council membership at the March 1985 Legislative Meeting.
4. Work actively with the National Center for Education Statistics (NCES) to ensure that reporting of data from all sources is accurate and timely.
5. Work with the major test publishers, survey organizations, corporations, and other groups that have large-scale assessment programs, to examine methods for developing state-by-state comparative data that are valid.
6. Work to develop interim measures of institutional improvement to meet the public need for short-term data on the effects of recent educational reforms. Examples of such measures are:
 - a. the increase in the number of academic courses taken by students;
 - b. the increase in the number of science teachers taking summer training; and
 - c. the increase in the number of students in dropout prevention programs.

To the U.S. Department of Education (ED)

1. Integrate the projects through which the Secretary reports on the state of Education in the nation by combining reports—The Condition of Education, Education Indicators Project, "Wall Chart," and other statistical and assessment reports. Request increased appropriations for the Department of Education (National Center for Education Statistics) assessment and evaluation efforts by five to six times current level (currently \$8+ million per year) to make the capability comparable with national reporting in health, agriculture, and other functions.

2. Over the short-run, in cooperation with the CCSSO, revise "State Education Statistics: State Performance Outcomes, Resource Inputs, and Population Characteristics, 1972, and 1982" prepared by the U.S. Department of Education, and released in January 1984 (popularly known as the Secretary's "Wall Chart") so that it: a) focuses on nationwide indices and trends; b) represents any state performance or outcome through trends or net change for the state itself rather than by rank order of states; c) includes Scholastic Aptitude Test (SAT) or American College Entrance Examination (ACT) scores only if reported along with the state percentages of twelfth graders represented by the SAT/ACT scores; d) includes student retention rates only if adjusted by net in/out migration and if data from different states are comparable; and e) includes the number of GED's awarded in the state and the change over the years. (It is important to recognize that post hoc analyses can never adequately adjust measures to ensure comparability. If comparisons are desired, they must be carefully planned in advance so that the indicators can be defined and weighted according to the questions they are being asked to answer.)
3. Integrate education assessment and evaluation results and other pertinent information from other federal agencies such as the Department of Defense (i.e., entrance test results), Department of Labor, National Science Foundation, and Census Bureau into the Education Department assessment and evaluation program.
4. Provide financial aid to states to stimulate new state assessment programs and strengthen existing state assessment efforts. Federal aid might also be used to relate state programs to the National Assessment of Educational Progress (NAEP) or other nationwide testing programs.

To the CCSSO and U.S. Department of Education Together

1. Continue joint federal/state/local planning efforts on assessment and evaluation. The efforts should include indicator model design and identification and use of joint opportunities for data gathering, analysis, and reporting. This includes examination of indices in other fields such as health, economics (CPI, Dow Jones Average), and welfare; and the development of indices for student achievement and for the relative challenge presented by differing groups of students (e.g., different socio-economic groupings) to meet the same achievement levels.
2. Establish a long-term commitment to develop a core set of education indicators and an accompanying analytical model which accurately portrays the educational system and effectively measures educational effectiveness. This effort will require two or three years, at least.
3. Educate the public, policy-makers, educators, and the media on the appropriate uses of data, assessment results, and education indicators, and the inferences that may legitimately be drawn from them.

4. Develop assessment and evaluation exchange programs with other nations and support United States participation in international or cross-national education assessment projects. Consider systematic support of the International Educational Achievement (IEA) effort. Also examine opportunities to work with the Organization of Economic and Community Development (OECD).

To States: Recommendations Regarding Statewide Indicators

1. Each state should develop a systematic approach to monitoring the status of education in relationship to state goals and context. Specifically:
 - a. A set of outcomes for the educational system and a way of measuring them should be agreed upon by various interests within each state.
 - b. Each state should develop a conceptual model which links outcomes to indicators critical to the quality of education. One approach to defining input indicators is to break them into four categories: 1) measures of the quality and level of resources such as teachers and curriculum materials; 2) measures of the processes within the schools such as time spent on instruction, teaching strategies, and in-service training approaches; 3) measures of the characteristics and needs of the students such as the percentage of handicapped and limited English-speaking ability; and 4) measures of the policy environment including the rules, regulations, and standards governing the state's educational system. The measures selected as input indicators should all have a solid and stable relationship to one or more of the outcome indicators.
 - c. Each state should publish a yearly report which contains displays and analyses of cross-sectional trend data showing the condition of its educational system as represented by outcomes and other indicators.
2. Each state should establish high standards to be followed in its collection and presentation of outcome indicators:
 - a. Multiple indicators should be presented (e.g., achievement scores, dropout rates, numbers of high school equivalencies, college-going rate, percentage of 18-19 year old youth in jobs and in the armed forces).
 - b. Data should be presented over time as trend and gain scores.
 - c. Work should be carried out to develop a "challenge index" which would give an indication of the gain of the state relative to its starting point.
 - d. Presentation of data should include ethnic, socio-economic, language dominance, and handicapped groupings.

To States: Recommendations Regarding Indicators Which Can be Compared Across States

Implementation of the following five recommendations shall be in accordance with a plan prepared by the CEIR Committee and presented to the Council at the Annual Meeting November 1985 for approval.

BEST COPY

1. The various states also should work with each other and with the CCSSO to establish a "core set" of input and outcome indicators to be collected on a yearly basis by each state. This would facilitate the examination of the effects of state policies by making comparisons available on a state-to-state and cluster-of-states basis.
 - a. The format and definition for each on the indicators would be common to all the states.
 - b. These indicators would not necessarily replace existing data gathered by the state; rather, they should be viewed as augmenting an existing data base.
 - c. The CCSSO and the states should work to make the indicators as parsimonious and inexpensive to collect as possible.
 - d. The indicators should be published by each state along with its other indicators in a yearly status report.
 - e. These indicators could be collected on a sample basis within states rather than for the total population concerned. This would decrease cost and burden and would make it clear that the indicators were not replacing existing data collection efforts.

(The implementation of this recommendation would take place over a two-year period and would be coordinated by the CCSSO.)

2. States should work with each other and with the CCSSO to establish a "core set" of outcome indicators. There should be multiple indicators including data from educational achievement tests. The CCSSO will sponsor a planning study to develop specifications and criteria for defining these outcome indicators.
3. With respect to achievement testing, the planning study should develop criteria for states to use in their selection of instruments for assessing student achievement. Over a period of approximately five years, the implementation of these criteria would result in states testing a common set of grades and a common set of skill and content areas with a set of tests that would be equatable across the states.
4. In addition, the achievement tests meeting the criteria should have a variety of special characteristics. They should: a.) assess basic and higher order conceptual skills; b.) assess a core of content knowledge in science, social studies, and English as well as in mathematics and reading; c.) provide the opportunity for a writing sample and a "production" item in some domain other than English; and d.) be of the highest possible quality and be continually upgraded as knowledge improves.

Given the current level of resource allocation and technical expertise available, developing and maintaining high quality testing instruments will be difficult. States will need to augment significantly the resources devoted to testing if this goal is to be met.

5. The "core" test(s) should augment rather than replace the existing state assessments of achievement. The primary function of the "core" tests would be to provide data representative of a state's progress and level of educational performance. Therefore, the test may be administered to samples rather than total populations of students at the selected grade levels. Samples per grade level can be quite small and provide reliable state level information. In addition there are a number of legitimate psychometric strategies which may be used to limit the testing time for any individual student. It would be straightforward to design a procedure that tested approximately 400 classrooms, each for one hour apiece, at each of three grade levels in a state and to come up with very reliable and extensive data. This would not generate data representative at the local level. Local districts could be given the option to augment the data collection to establish their own representative sample.

To Local Education Agencies

We acknowledge that recommendations made to local districts should be given more attention in follow-up documents. Much of the burden in the proposed activities will fall to them, and the recommendations directed toward them could be significantly enhanced.

1. The CCSSO encourages local education agencies (LEAs) to make effective use of diagnostic and evaluative instruments to improve the quality of instruction in their schools. The CCSSO will hold a conference on the use of diagnostic test instruments which will assess the state of the art and, it is hoped, provide useful information for local districts.
3. The CCSSO encourage local districts to establish systematic data collection approaches in order to provide timely and accurate data to concerned citizens and policy makers. These data should include outcome indicators (such as achievement, retention, and college-going rates) as well as input indicators of the sort suggested for the states.